Remarks

Amendments to the Claims

Each of independent claims 1, 12, 33, 52, and 59 is amended to recite that the biological subject is "obtained" from a region of the mammal. Similarly, each of independent claims 9, 22, and 62 is amended to recite that the second sample is "obtained" from the patient. The specification supports these amendments on page 22, line 11 to page 23, line 2, where it describes obtaining biological subjects and samples.

Each of independent claims 9, 22, and 62 is amended to clarify that the recited steps are carried out at two different time points during a therapeutic regimen. The specification supports this amendment on page 56, lines 13-23, which describes taking samples from patients "at various time points during the course of the application of a treatment regimen."

New claims 65 and 66 recite that the mammal or patient, respectively, is a human, and new claims 67-82 recite human hepsin sequences SEQ ID NO:1 (nucleotide) and SEQ ID NO:2 (amino acid), which are disclosed in the specification as filed.

The amendments do not add new matter.

The Rejection of Claims 59-64 Under 35 U.S.C. § 112 ¶ 2

The Office Action maintains the rejection of claims 59-64 under 35 U.S.C. § 112 ¶ 2 because the recitations "first indirect measure" and "second indirect measure" allegedly are indefinite. Applicants respectfully traverse the rejection.

Under the second paragraph of 35 U.S.C. § 112, the relevant inquiry

> . . . is merely to determine whether the claims do, in fact, set out and circumscribe a particular area with a reasonable degree of precision and particularity. It is here where the definiteness of the

> language employed must be analyzed -- not in a vacuum, but always in light of the teachings of the prior art and of the particular application disclosure as it would be interpreted by one possessing the ordinary level of skill in the pertinent art.

In re Moore, 439 F.2d 1232, 1235, 58 C.C.P.A. 1042, 1046-47 (1971) (footnote omitted). The Office Action contends that "[t]he terms, 'first indirect measure' or 'second indirect measure' are not defined by the claim(s), the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the metes and bounds of such terms." Office Action at page 3. On the contrary, one of ordinary skill in the art would understand that, in the context of determining gene copy number, any method other than directly counting copies of genes is, by definition, an indirect measure of gene copy number. Indirect methods of detecting gene copy number were well known in the art when this application was filed. As the specification teaches, these methods include Southern blotting, *in situ* hybridization, comparative genomic hybridization (CGH), amplification-based assays (*e.g.*, a PCR-based TaqMan assay), and DNA microarray-based CGH (pages 41-43; Example 1). Each of these methods detects a measurement which reflects gene copy number and is, therefore, indirect. In the context of the specification and the knowledge of one of ordinary skill, the term "indirect measure" as used in claims 59-64 is definite.

Under 35 U.S.C. § 112, second paragraph, the claims must "reasonably apprise those skilled in the art both of the utilization and scope of the invention." *Georgia-Pacific Corp. v. United States Plywood Corp.*, 258 F.2d 124, 136, 118 U.S.P.Q. 122, 130 (2d Cir. 1958), *cert. denied*, 358 U.S. 884 (1958). Claims 59-64 serve this purpose. Claims 59-64 are therefore definite.

Applicants respectfully request withdrawal of the rejection.

Rejection of Claims 1-3, 9-12, 22-24, 33-35, and 39-64 Under 35 U.S.C. § 112 ¶ 1

Claims 1-3, 9-12, 22-24, 33-35, and 39-64 stand rejected under 35 U.S.C. § 112 ¶ 1 as insufficiently described. The Office Action asserts the term "hepsin" is broader than the description supports. Applicants respectfully traverse the rejection.

What is required to satisfy the written description requirement depends on the nature of the invention claimed. *In re DiLeone*, 436 F.2d 1404, 1405, 168 U.S.P.Q. 592, 593 (C.C.P.A. 1971). Independent claims 1, 52, and 59 recite a hepsin gene copy number in a biological subject obtained from a region of a mammal that is suspected to be precancerous or cancerous. Independent claims 9 and 62 recite a hepsin gene copy number in a sample of precancerous of cancer cells obtained from a patient. Independent claim 12 recites a test level of hepsin mRNA expression in a biological subject obtained from a region of the mammal that is suspected to be precancerous or cancerous. Independent claim 22 recites expression levels of at least one of hepsin mRNA or hepsin protein in samples of a biological subject comprising precancerous or cancer cells obtained from a patient. Independent claim 33 recites detecting a test hepsin protein expression level in a biological subject obtained from a region of the mammal that is suspected to be precancerous or cancerous. In each case, the recited hepsin gene, protein, or mRNA is one which naturally occurs in the mammal.

The naturally occurring mammalian hepsin protein and gene were well known in the art at the February 14, 2001 priority date of this application. *See, e.g.*, Kawamura[1] and Vu[2] (mouse

[1] Kawamura *et al.*, "Complete nucleotide sequence, origin of isoform and functional characterization of the mouse hepsin gene," Eur J Biochem. 1999 Jun;262(3):755-64.

[2] Vu *et al.*, "Identification and Cloning of the Membrane-associated Serine Protease, Hepsin, from Mouse Preimplantation Embryos," *J. Biol. Chem. 272*, 31315-20, 1997.

hepsin); Kazama[3] and Leytus[4] (human hepsin); and Farley (rat hepsin),[5] each provided with the accompanying Information Disclosure Statement. An adequate written description of a gene which is well known in the art does not require a structural recitation either in the specification or in the claims. *See Capon v. Eshhar*, 418 F.3d 1349, 1360-61, 76 U.S.P.Q.2d 1078, 1087 (Fed. Cir. 2005) ("the Board erred in ruling that § 112 imposes a *per se* rule requiring recitation in the specification of the nucleotide sequence of claimed DNA, when that sequence is already known in the field."). Thus, the fact that the claims do not recite a sequence identifier "relating to hepsin" as noted in the Office Action at page 4 ¶ 1 does not mean that the claims lack written description. Under *Capon*, a sequence identifier is not required to describe the term "hepsin."

The Office Action cites nine different GenBank Accession Numbers as disclosing hepsin genes. Page 4 ¶ 1. The hepsin coding sequences with these Accession Numbers, however, are not "highly variant," as the Office Action contends. *Id.*

Accession Numbers NM_017112 and X70900 are rat hepsin coding sequences; a BLAST alignment of these sequences shows that the two sequences are 100% identical (Exhibit 1). Accession Numbers NM_008281 and AF030065 are mouse hepsin coding sequences; a BLAST alignment of these sequences shows that the two sequences are 100% identical (Exhibit 2). The mouse and rat sequences are 94% identical (Exhibit 3).

Accession Numbers BC025716.1, NM_182983.1, NM_002151.1, X07732, and X07002.1 are human hepsin coding sequences. BLAST alignments of SEQ ID NO:1 with each of the other

---

[3] Kazama *et al.*, "Hepsin, a putative membrane-associated serine protease, activates human factor VII and initiates a pathway of blood coagulation on the cell surface leading to thrombin formation," J Biol Chem. 1995 Jan 6;270(1):66-72.

[4] Leytus *et al.*, "A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells," Biochemistry 27, 1067-1074, 1988.

[5] Farley *et al.*, "Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase," Biochim Biophys Acta. 1993 Jun 25;1173(3):350-2.

human hepsin coding sequences are provided in Exhibits 4-8. SEQ ID NO:1 is 100% identical over 1593 contiguous nucleotides and 100% identical over 193 contiguous nucleotides with each of splice variants NM_182983.1 (Exhibit 5) and X07732 (Exhibit 7). SEQ ID NO:1 is 99% identical with BC025716.1 (Exhibit 4), 100% identical with NM_002151.1 (Exhibit 6), and 100% identical with 1199 contiguous nucleotides of X07002.1 (Exhibit 8). SEQ ID NO:1 is 83% identical to a long stretch (1318 nucleotides) of the rat and mouse hepsin coding sequences (Exhibit 9).

Mammalian hepsin coding sequences were known in the art when this application was filed, and none of the sequences cited in the Office Action are "highly variant." The specification therefore satisfies the written description requirement with respect to the term "hepsin" as recited in the rejected claims.

Applicants respectfully request withdrawal of the rejection.

Enablement Rejections

All of the pending claims stand rejected under 35 U.S.C. § 112 ¶ 1 as not enabled for their full scope. In each case, the Office Action contends that the specification enables the claimed methods for *ex vivo* use but not for use *in vivo*. Office Action at page 7 ¶ 1. Applicants respectfully traverse the rejection.

To advance prosecution, each of the independent claims has been amended to clarify that the recited method steps are carried out *ex vivo*. Each of the independent diagnostic method claims (1, 12, 33, 52, and 59) now recites that the biological subject is "obtained from" a region of the mammal suspected to be precancerous or cancerous. Similarly, each of the independent methods of monitoring treatment efficacy (9, 22, and 62) recites that the sample is "obtained

from" a patient. Because the Office Action acknowledges this subject matter is enabled, Applicants respectfully request withdrawal of the rejection.

Respectfully submitted,

BANNER & WITCOFF, LTD.

/Lisa M. Hemmendinger/

Date: October 27, 2006 By: _____

Lisa M. Hemmendinger
Registration No. 42,653

Customer No. 22907

Exhibit 1. BLAST alignment of NM_017112 and X70900

NCBI **Blast 2 Sequences results**

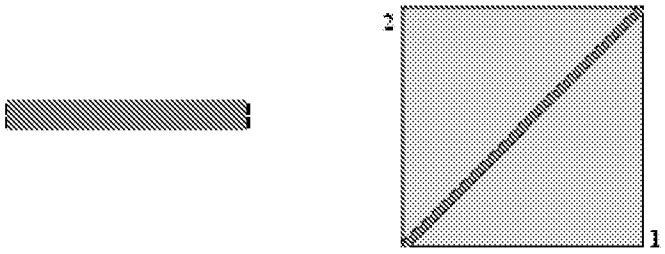
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 30 expect: 10.0000 wordsize: 11 Filter: ☐ View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: gi|8393559|ref|NM_017112.1|Rattus norvegicus hepsin (Hpn), mRNA
Length = 1739 (1 .. 1739)

Sequence 2: gi|57928|emb|X70900.1|RNHEPAR.norvegicus mRNA for hepsin
Length = 1739 (1 .. 1739)



NOTE:Bitscore and expect value are calculated based on the size of the nr database.

NOTE:If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3344 bits (1739),  Expect = 0.0
 Identities = 1739/1739 (100%), Gaps = 0/1739 (0%)
 Strand=Plus/Plus

Query  1     GCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTCCCGGGCTGCCCGCTGCTGT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     GCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTCCCGGGCTGCCCGCTGCTGT  60

Query  61    GGGGACACCATGCCCCTTCCCAGGCCCGGAGACTAACCCCAAACCTGCACCATCTCCGGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGGGACACCATGCCCCTTCCCAGGCCCGGAGACTAACCCCAAACCTGCACCATCTCCGGC  120

Query  121   GAACCCCAGGGTTCCGCCCCAGCCCAACAGGTCAACCTGGGAATCATTAACAAGAGTCCC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1021  TGCGGGACAGGCCCTGGTGGACGGCAAGGTCTGTACAGTGACCGGCTGGGGTAACACACA  1080

Query  1081  GTTCTATGGCCAGCAAGCTGTGGTGCTCCAAGAGGCCCGGGTCCCCATCATAAGCAACGA  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  GTTCTATGGCCAGCAAGCTGTGGTGCTCCAAGAGGCCCGGGTCCCCATCATAAGCAACGA  1140

Query  1141  AGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAACCCAAGATGTTCTGTGCTGG  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  AGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAACCCAAGATGTTCTGTGCTGG  1200

Query  1201  CTATCCTGAGGGTGGTATTGATGCATGCCAGGGTGACAGCGGAGGCCACTTTGTATGTGA  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  CTATCCTGAGGGTGGTATTGATGCATGCCAGGGTGACAGCGGAGGCCACTTTGTATGTGA  1260

Query  1261  GGACAGAATCTCTGGAACATCAAGATGGCGGCTCTGCGGCATTGTAAGCTGGGGTACGGG  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  GGACAGAATCTCTGGAACATCAAGATGGCGGCTCTGCGGCATTGTAAGCTGGGGTACGGG  1320

Query  1321  CTGTGCTTTGGCCCGGAAGCCGGGAGTGTACACCAAAGTCATTGACTTCCGGGAGTGGAT  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  CTGTGCTTTGGCCCGGAAGCCGGGAGTGTACACCAAAGTCATTGACTTCCGGGAGTGGAT  1380

Query  1381  CTTCCAGGCCATAAAGACTCACTCCGAAGCTACCGGCATGGTAACTCAGCCCTGACCCCG  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  CTTCCAGGCCATAAAGACTCACTCCGAAGCTACCGGCATGGTAACTCAGCCCTGACCCCG  1440

Query  1441  CCTCATCGGCCTGCTCCGCGCTGCTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGGCTCCA  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  CCTCATCGGCCTGCTCCGCGCTGCTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGGCTCCA  1500

Query  1501  GCCGCACGTGGCAGGCTCCACACTGGGCCTCACATGGAACGGTTTTCTGCTCGGATCCAG  1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  GCCGCACGTGGCAGGCTCCACACTGGGCCTCACATGGAACGGTTTTCTGCTCGGATCCAG  1560

Query  1561  TCCATAGATCCAAGGATGCTGGGTCCAAGGACCTCTCTTCCACAGTGGCCGGCCCACTCA  1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  TCCATAGATCCAAGGATGCTGGGTCCAAGGACCTCTCTTCCACAGTGGCCGGCCCACTCA  1620

Query  1621  ATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAAATATTACTCTGTCCTCTGGGG  1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  ATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAAATATTACTCTGTCCTCTGGGG  1680

Query  1681  GCTGCTTTCGAGGCGCCCCTTGTGCGGATGCTCTTTAAATAATAAAGGTGGTTTTGATT  1739
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  GCTGCTTTCGAGGCGCCCCTTGTGCGGATGCTCTTTAAATAATAAAGGTGGTTTTGATT  1739



CPU time:     0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
```

```
Number of Sequences: 1
Number of Hits to DB: 482
Number of extensions: 1
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1739
Length of database: 17,904,826,723
Length adjustment: 27
Effective length of query: 1712
Effective length of database: 17,904,826,701
Effective search space: 30653063313824
Effective search space used: 30653063313824
X1: 11 (21.1 bits)
X2: 15 (28.8 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 2.

Blast 2 Sequences results

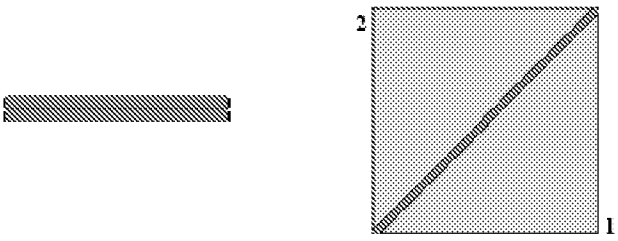
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter ☐ View option Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: gi|2606036|gb|AF030065.1|AF030065
Length = 1781 (1 .. 1781)

Sequence 2: gi|6680266|ref|NM_008281.1|Mus musculus hepsin (Hpn), mRNA
Length = 1781 (1 .. 1781)



NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3424 bits (1781),  Expect = 0.0
 Identities = 1781/1781 (100%), Gaps = 0/1781 (0%)
 Strand=Plus/Plus

Query  1     CGGAGGAGGGGGCTCGGCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTTCCG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     CGGAGGAGGGGGCTCGGCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTTCCG  60

Query  61    GGCTGTCCGCTGCTGTGGGGACAGACACCATGCCCCTGCCCAGGCCCGGAGACTAACCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGCTGTCCGCTGCTGTGGGGACAGACACCATGCCCCTGCCCAGGCCCGGAGACTAACCCC  120

Query  121   AAACCTGCACCATCTCCGGCGAACCCCAGGGTTCCGCCCCAGCCCAACAGGTCAACCTGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  121   AAACCTGCACCATCTCCGGCGAACCCCAGGGTTCCGCCCCAGCCCAACAGGTCAACCTGG  180

Query  181   GAATCATTAACAAGAGTCCCTGACATGGCGAAGGAGGGTGGCCGGACTGCAGCATGCTGC  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   GAATCATTAACAAGAGTCCCTGACATGGCGAAGGAGGGTGGCCGGACTGCAGCATGCTGC  240

Query  241   TCCAGACCCAAGGTGGCAGCTCTCATTGTGGGTACCCTGCTGTTCCTGACAGGCATTGGG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   TCCAGACCCAAGGTGGCAGCTCTCATTGTGGGTACCCTGCTGTTCCTGACAGGCATTGGG  300

Query  301   GCCGCGTCCTGGGCCATTGTGACCATCCTACTGCAGAGTGACCAGGAGCCACTGTACCAA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   GCCGCGTCCTGGGCCATTGTGACCATCCTACTGCAGAGTGACCAGGAGCCACTGTACCAA  360

Query  361   GTGCAGCTCAGTCCAGGGGACTCACGACTTGCAGTGTTGGACAAGACGGAGGGTACGTGG  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   GTGCAGCTCAGTCCAGGGGACTCACGACTTGCAGTGTTGGACAAGACGGAGGGTACGTGG  420

Query  421   AGGCTACTGTGCTCCTCACGCTCCAATGCCAGGGTGGCAGGGCTCGGCTGTGAGGAGATG  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   AGGCTACTGTGCTCCTCACGCTCCAATGCCAGGGTGGCAGGGCTCGGCTGTGAGGAGATG  480

Query  481   GGCTTTCTCAGGGCTCTGGCGCACTCGGAGCTGGATGTGCGCACTGCGGGCGCCAACGGC  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   GGCTTTCTCAGGGCTCTGGCGCACTCGGAGCTGGATGTGCGCACTGCGGGCGCCAACGGC  540

Query  541   ACATCGGGCTTCTTTTGCGTGGACGAGGGCGGACTGCCTCTGGCTCAGAGGTTGCTGGAT  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   ACATCGGGCTTCTTTTGCGTGGACGAGGGCGGACTGCCTCTGGCTCAGAGGTTGCTGGAT  600

Query  601   GTCATCTCTGTATGTGACTGTCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGT  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   GTCATCTCTGTATGTGACTGTCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGT  660

Query  661   GGCCGCAGGAAGCTGCCGGTGGACCGCATTGTGGGGGGCCAGGACAGCAGTCTGGGAAGG  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   GGCCGCAGGAAGCTGCCGGTGGACCGCATTGTGGGGGGCCAGGACAGCAGTCTGGGAAGG  720

Query  721   TGGCCGTGGCAGGTCAGCCTGCGTTATGATGGGACCCACCTCTGTGGGGGGTCCCTGCTG  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   TGGCCGTGGCAGGTCAGCCTGCGTTATGATGGGACCCACCTCTGTGGGGGGTCCCTGCTG  780

Query  781   TCTGGGGACTGGGTGCTGACTGCTGCACATTGCTTTCCAGAGCGGAACCGGGTCCTGTCT  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   TCTGGGGACTGGGTGCTGACTGCTGCACATTGCTTTCCAGAGCGGAACCGGGTCCTGTCT  840

Query  841   CGGTGGCGAGTATTTGCTGGTGCTGTAGCCCGGACCTCACCCCATGCTGTGCAACTGGGG  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   CGGTGGCGAGTATTTGCTGGTGCTGTAGCCCGGACCTCACCCCATGCTGTGCAACTGGGG  900

Query  901   GTTCAGGCTGTGATCTATCATGGGGGCTACCTTCCCTTTCGAGACCCTACTATTGACGAA  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   GTTCAGGCTGTGATCTATCATGGGGGCTACCTTCCCTTTCGAGACCCTACTATTGACGAA  960

Query  961   AACAGCAATGACATTGCCTTGGTCCACCTCTCTAGCTCCCTGCCTCTCACAGAATACATC  1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   AACAGCAATGACATTGCCTTGGTCCACCTCTCTAGCTCCCTGCCTCTCACAGAATACATC  1020

Query  1021  CAGCCAGTGTGTCTCCCTGCTGCGGGACAGGCCCTGGTGGATGGCAAGGTCTGTACTGTG  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1021  CAGCCAGTGTGTCTCCCTGCTGCGGGACAGGCCCTGGTGGATGGCAAGGTCTGTACTGTG  1080

Query  1081  ACCGGCTGGGGTAACACACAGTTCTATGGCCAACAGGCTATGGTGCTCCAAGAGGCCCGG  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  ACCGGCTGGGGTAACACACAGTTCTATGGCCAACAGGCTATGGTGCTCCAAGAGGCCCGG  1140

Query  1141  GTTCCCATCATAAGCAACGAAGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAG  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  GTTCCCATCATAAGCAACGAAGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAG  1200

Query  1201  CCCAAGATGTTCTGTGCTGGCTATCCTGAGGGTGGCATTGATGCGTGCCAGGGCGACAGT  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  CCCAAGATGTTCTGTGCTGGCTATCCTGAGGGTGGCATTGATGCGTGCCAGGGCGACAGT  1260

Query  1261  GGAGGCCCCTTTGTGTGTGAAGACAGCATCTCTGGGACATCAAGGTGGCGGCTATGTGGC  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  GGAGGCCCCTTTGTGTGTGAAGACAGCATCTCTGGGACATCAAGGTGGCGGCTATGTGGC  1320

Query  1321  ATTGTAAGCTGGGGTACCGGCTGTGCTTTGGCCCGGAAGCCAGGAGTGTACACCAAAGTC  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  ATTGTAAGCTGGGGTACCGGCTGTGCTTTGGCCCGGAAGCCAGGAGTGTACACCAAAGTC  1380

Query  1381  ACTGACTTTCCGGGAGTGGATCTTCAAGGCCATAAAGACTCACTCCGAAGCCAGTGGCATG  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  ACTGACTTTCCGGGAGTGGATCTTCAAGGCCATAAAGACTCACTCCGAAGCCAGTGGCATG  1440

Query  1441  GTGACTCAGCCCTGATCCCGCCTCATCTCGCTGCTCCGTGCTGCACTAGCATCCAGAGTC  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  GTGACTCAGCCCTGATCCCGCCTCATCTCGCTGCTCCGTGCTGCACTAGCATCCAGAGTC  1500

Query  1501  AGAGTTGGTCTGGTGGCTCCAGCCCCACGTGGTAGGCTCCACACTGGGCCTCACATGGAA  1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  AGAGTTGGTCTGGTGGCTCCAGCCCCACGTGGTAGGCTCCACACTGGGCCTCACATGGAA  1560

Query  1561  TGGTTTCCTGCTCAGATCCAGTCCACGGGTCCAAGGATGCTGGATCCAAGGACTTCTCTT  1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  TGGTTTCCTGCTCAGATCCAGTCCACGGGTCCAAGGATGCTGGATCCAAGGACTTCTCTT  1620

Query  1621  CCACAGTGGCCGGCCCACTCAATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAA  1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  CCACAGTGGCCGGCCCACTCAATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAA  1680

Query  1681  ATATTACTCTGTCCTCTGGGGGGCGCTCTAGGGAGCCCCTTGTGCAGATGCTCTTTAAAT  1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  ATATTACTCTGTCCTCTGGGGGGCGCTCTAGGGAGCCCCTTGTGCAGATGCTCTTTAAAT  1740

Query  1741  AATAAAGGTGGTTTTGATTAATGGGAGAAAAAAAAAAAAAA  1781
              ||||||||||||||||||||||||||||||||||||||||
Sbjct  1741  AATAAAGGTGGTTTTGATTAATGGGAGAAAAAAAAAAAAAA  1781
```

```
CPU time:     0.02 user secs.       0.01 sys. secs          0.03 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


 Matrix: blastn matrix:1 -2
 Gap Penalties: Existence: 5, Extension: 2
 Number of Sequences: 1
 Number of Hits to DB: 517
 Number of extensions: 10
 Number of successful extensions: 1
 Number of sequences better than 10.0: 1
 Number of HSP's gapped: 1
 Number of HSP's successfully gapped: 1
 Length of query: 1781
 Length of database: 17,904,826,729
 Length adjustment: 27
 Effective length of query: 1754
 Effective length of database: 17,904,826,702
 Effective search space: 31405068035308
 Effective search space used: 31405068035308
 X1: 11 (21.1 bits)
 X2: 26 (50.0 bits)
 X3: 26 (50.0 bits)
 S1: 14 (27.6 bits)
 S2: 22 (43.0 bits)
```

Exhibit 3. BLAST alignment of rat and mouse hepsin coding sequences

NCBI **Blast 2 Sequences results**

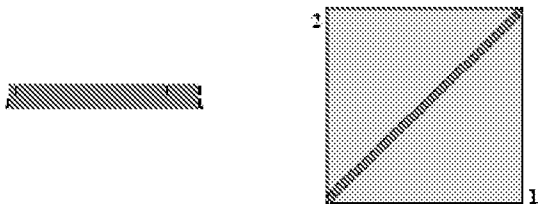
Pubmed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter View option Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: gi|8393559|ref|NM_017112.1|Rattus norvegicus hepsin (Hpn), mRNA
Length = 1739 (1 .. 1739)

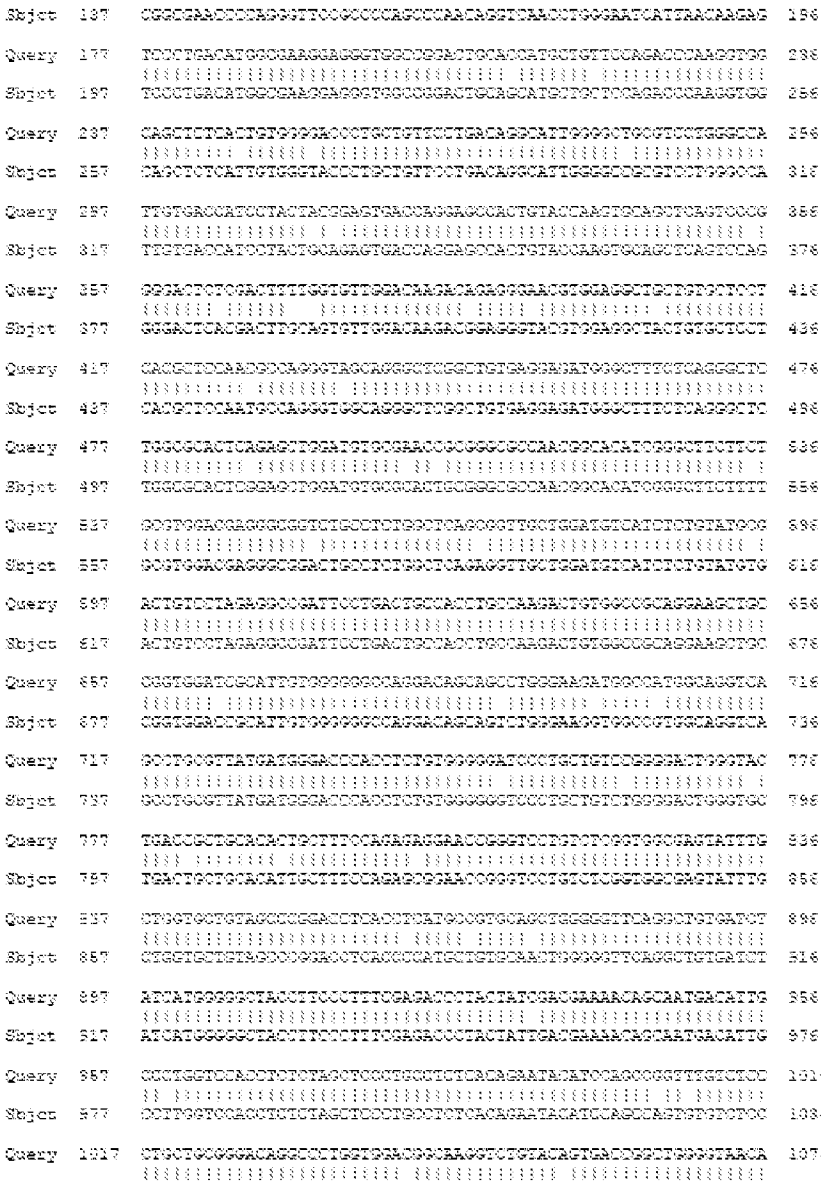
Sequence 2: gi|6680266|ref|NM_008281.1|Mus musculus hepsin (Hpn), mRNA
Length = 1781 (1 .. 1781)

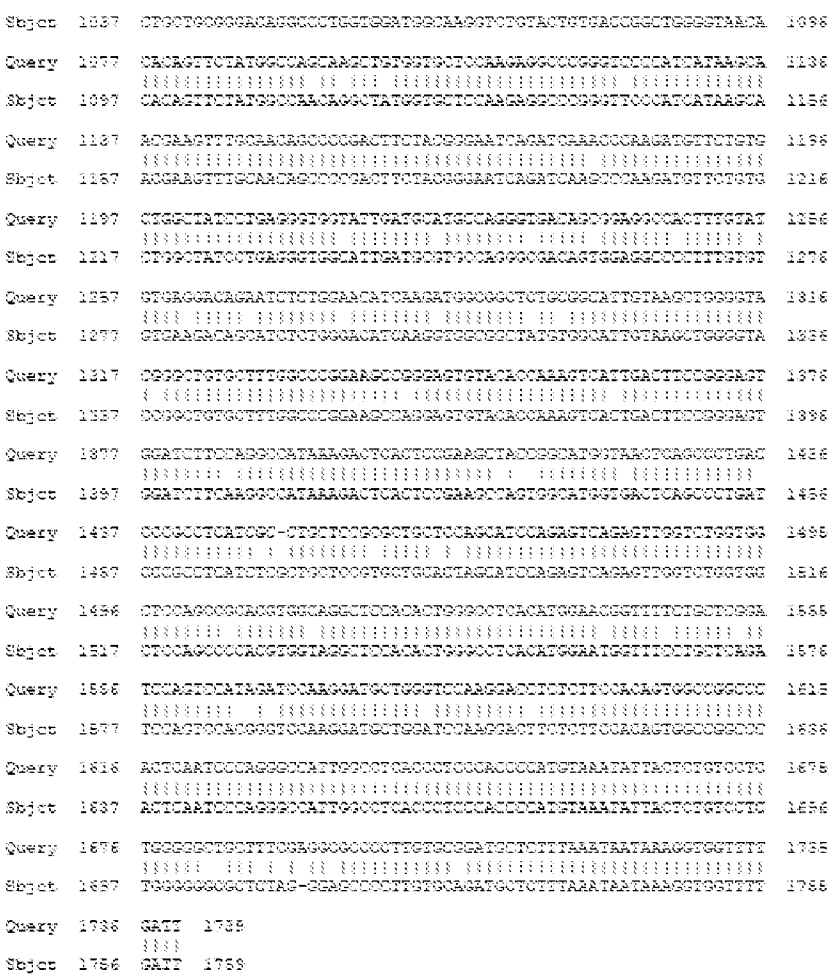


NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 2789 bits (1438),  Expect = 0.0
 Identities = 1646/1744 (94%), Gaps = 6/1744 (0%)
 Strand=Plus/Plus

Query  1     GCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCCCCCGGGGCTGCCCGCTGCTGT  60
             ||||||||||||||||||||||||||||||||||||||||||| |||||||| ||||||||||
Sbjct  17    GCAGGCCCCACGCCGCTGGCTGCTGCTGCCACCCTTGCCTTCCGGGCTGCTGCCCGCTGCTGT  76

Query  61    GGGGACA----CCATGCCCCTTCCCAGGCCCGGAGACTAACCCCAAACCTGCACCATCTC  116
             |||||||    |||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  77    GGGGACAGACACCATGCCCCTGCCCAGGCCCGGAGACTAACCCCAAACCTGCACCATCTC  136

Query  117   CGGCGAAGCCCGAGGGTTCCGCCCTAGCCCAACAGGTCAACCTGGGAATCATTAACAAGAG  176
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```





```
CPU time:      0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


  Matrix: blastn matrix:1 -2
  Gap Penalties: Existence: 5, Extension: 2
  Number of Sequences: 1
  Number of Hits to DB: 331
  Number of extensions: 7
  Number of successful extensions: 4
  Number of sequences better than 10.0: 1
  Number of HSP's gapped: 1
  Number of HSP's successfully gapped: 1
  Length of query: 1739
  Length of database: 18,244,403,944
  Length adjustment: 27
  Effective length of query: 1712
  Effective length of database: 18,244,403,917
  Effective search space: 31234419505904
  Effective search space used: 31234419505904
  X1: 11 (21.1 bits)
  X2: 26 (50.0 bits)
  X3: 26 (50.0 bits)
  S1: 14 (27.6 bits)
  S2: 22 (43.0 bits)
```

Exhibit 4. BLAST alignment of SEQ ID NO:1 with BC025716.1

NCBI **Blast 2 Sequences results**

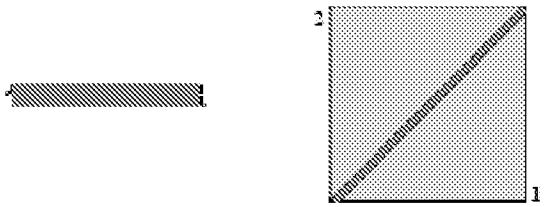
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|19343934|gb|BC025716.1|Homo sapiens hepsin (transmembrane protease, serine 1), mRNA (cDNA clone MGC:34588 IMAGE:5228525), complete cds
Length = 1761 (1 .. 1761)



NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3330 bits (1732),  Expect = 0.0
 Identities = 1734/1735 (99%), Gaps = 0/1735 (0%)
 Strand=Plus/Plus

Query  49    GCCTGGCCTAGCAGGCCCCACGCCACCGCCTCTGCCTCCAGGCCGCCCGCTGCTGCCCCCC  108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     GCCTGGCCTAGCAGGCCCCACGCCACCGCCTCTGCCTCCAGGCCGCCCGCTGCTGCCCCCC  60

Query  109   CCACCATGCTCCTGCCCAGGCCTGGAGACTGACCCGACCCCTGGCACTACCTCGAGGCTCC  168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    CCACCATGCTCCTGCCCAGGCCTGGAGACTGACCCGACCCCTGGCACTACCTCGAGGCTCC  120

Query  169   GCCCCCACCTGCTGGACCCCAGGGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATGAT  228
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121   GCCCCCACCTGCTGGACCCCAGGGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATCAT  180

Query  229   TAACAAGAGGCAGTGACATGGGGCAGAAGGAGGGTGGCCGGACTGTGCCATGCTGCTCCA  288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   TAACAAGAGGCAGTGACATGGCGCAGAAGGAGGGTGGCCGGACTGTGCCATGCTGCTCCA  240

Query  289   GACCCAAGGTGGCAGCTCTCACTGCGGGGACCCTGCTACTTCTGACAGCCATCGGGGCGG  348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   GACCCAAGGTGGCAGCTCTCACTGCGGGGACCCTGCTACTTCTGACAGCCATCGGGGCGG  300

Query  349   CATCCTGGGCCATTGTGGCTGTTCTGCTCAGGAGTGACCAGGAGCCGCTGTACCCAGTCC  408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   CATCCTGGGCCATTGTGGCTGTTCTGCTCAGGAGTGACCAGGAGCCGCTGTACCCAGTCC  360

Query  409   AGGTCAGCTCTGCGGACGCTCGGCTCATGGTCTTTGACAAGACGGAAGGGACGTGGCGGC  468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   AGGTCAGCTCTGCGGACGCTCGGCTCATGGTCTTTGACAAGACGGAAGGGACGTGGCGGC  420

Query  469   TGCTGTGCTCCTCGGGGCTCCAACGCCAGGGTAGCCGGACTCAGCTGCGAGGAGATGGGCT  528
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   TGCTGTGCTCCTCGCGCTCCAACGCCAGGGTAGCCGGACTCAGCTGCGAGGAGATGGGCT  480

Query  529   TCCTCAGGGCACTGACCCACTCCGAGCTGGAACGTGCGAACGGCGGGGCGCCAATGGCACGT  588
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   TCCTCAGGGCACTGACCCACTCCGAGCTGGAACGTGCGAACGGCGGGGCGCCAATGGCACGT  540

Query  589   CGGGGCTTCTTCTGTGTGGACGAGGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAGGTCA  648
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   CGGGCTTCTTCTGTGTGGACGAGGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAGGTCA  600

Query  649   TCTCCGTGTGTGATTGCCGCAGAGGCCGTTTCTTGGCCGCCATGTGCCAAGACTGTGCCC  708
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   TCTCCGTGTGTGATTGCCCCAGAGGCCGTTTCTTGGCCGCCATCTGCCAAGACTGTGCCC  660

Query  709   GCAGGAAGCTGCCCGTGGACCGCATCGTGGGAGGCCGGACACCAGCTTGGGCCGGTGGC  768
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   GCAGGAAGCTGCCCGTGGACCGCATCGTGGGAGGCCGGACACCAGCTTGGGCCGGTGGC  720

Query  769   CGTGGCAAGTCAGCCTTCGCTATGATGGAGCAGACCTGTGTGGGGGATCCCTGCTCTCCG  828
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   CGTGGCAAGTCAGCCTTCGCTATGATGGAGCAGACCTCTGTGGGGGATCCCTGCTCTCCG  780

Query  829   GGGACTGGGTGCTGACAGCCGCCACTGCTTCCCGGAGCGGAACCGGGTCTTGTCCCGAT  888
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  781   GGGACTGGGTGCTGACAGCCCCCATTGCTTCCCGGAGCGGAACCCGGGTCCTGTCCCGAT  840

Query  889   GGCGAGTGTTTGCCGGTGCCGTGGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGGGTCC  948
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   GGCGAGTGTTTGCCGGTGCCGTGGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGGGTGC  900

Query  949   AGGCTGTGGTCTACCACGGGGGCTATGTTCCGCTTTCGGGACCCCAACACGCGAGGAGAACA  1008
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   AGGCTGTGGTCTACCACGGGGGCTATGTTCCGCTTTCGGGACCCCAACACGCGAGGAGAACA  960

Query  1009  GCAACGATATTGCCCTGGTCCACCTCTCCAGTCCCCTGCCGCTCACAGAATACATCCAGC  1068
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   GCAACGATATTGCCCTGGTCCACCTCTCCAGTCCCCTGCCGCTCACAGAATACATCCAGC  1020

Query  1069  CTGTGTGCCTTCCAGTTGCCGGCCAGGCCCTGGTGGAIGGCAAGAICTGTACCGTGACGG  1128
```

```
CPU time:     0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda      K      H
    1.33    0.621     1.12

Gapped
Lambda      K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2



Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 489
Number of extensions: 5
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1788
Length of database: 18,244,403,944
Length adjustment: 17
Effective length of query: 1756
Effective length of database: 18,244,403,817
Effective search space: 32037173273852
Effective search space used: 32037173273852
X1: 11 (21.1 bits)
X2: 25 (50.0 bits)
X3: 28 (56.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 5. BLAST alignment of SEQ ID NO:1 with NM_182983.1

NCBI **Blast 2 Sequences results**

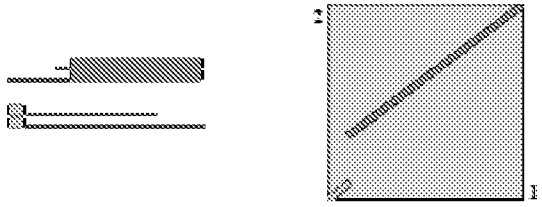
Publied Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|33695154|ref|NM_182983.1|Homo sapiens hepsin (transmembrane protease, serine 1) (HPN), transcript variant 1, mRNA
Length = 2363 (1 .. 2363)



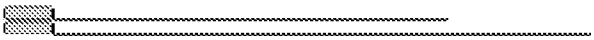
NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3053 bits (1653),  Expect = 0.0
 Identities = 1653/1653 (100%), Gaps = 0/1653 (0%)
 Strand=Plus/Plus

Query  191   GGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATCATTAACAAGAGGCAATGGACATGGC  250
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  771   GGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATCATTAACAAGAGGCAATGGACATGGC  830

Query  251   GCAGAAGGAGGGTGGCCGGACTGTGCCATGCTGCTCCAGACCCAAGGTGGCAGCTCTCAC  310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  831   GCAGAAGGAGGGTGGCCGGACTGTGCCATGCTGCTCCAGACCCAAGGTGGCAGCTCTCAC  890

Query  311   TGCGGGGACCCTGCTACTTCTGACAGCCATCGGGGCGGCATCCTGGGCCATTGTGGCTGT  370
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881   TGCGGGGACCCTGCTACTTCTGACAGCCATCGGGGCGGCATCCTGGGCCATTGTGGCTGT  940

Query  371   TCTCCTCAGGAGTGACCAGGAGCCGCTGTACCCAGTGCAGGTCAGCTCTGCGGACGCTCG  430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  941   TCTCCTCAGGAGTGACCAGGAGCCGCTGTACCCAGTGCAGGTCAGCTCTGCGGACGCTCG 1000

Query  431   GCTCATGGTCTTTGACAAGACGGAAGGGACGTGGCGGCTGTGTGCTCCTGGGCTCCAA  490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1001  GCTCATGGTCTTTGACAAGACGGAAGGGACGTGGCGGCTGTGTGCTCCTGGGCTCCAA 1060

Query  491   CGCCAGGGTAGCCGGACTCAGCTGCGAGGAGATGGGGTTCCTCAGGGCACTGACCCACTC  550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1061  CGCCAGGGTAGCCGGACTCAGCTGCGAGGAGATGGGGTTCCTCAGGGCACTGACCCACTC 1120

Query  551   CGAGCTGGACGTGCGAACGGGGGGGGCCAATGGCACGTCGGGGTTCTTCTGTGTGGACGA  610
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1121  CGAGCTGGACGTGCGAACGGGGGGGCCAATGGCACGTCGGGGTTCTTCTGTGTGGACGA 1180

Query  611   GGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAGGTCATCTCCGTGTGTGATTGCCCCAG  670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1181  GGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAGGTCATCTCCGTGTGTGATTGCCCCAG 1240

Query  671   AGGCCGTTTCTTGGGCGCCATCTGCCAAGACTGTGGCCGCAGGAAGCTGCCCGTGGACCG  730
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1241  AGGCCGTTTCTTGGCCGCCATCTGCCAAGACTGTGGCCGCAGGAAGCTGCCCGTGGACCG 1300

Query  731   CATCGTGGTAGGCCGGGACACCAGCTTGGGCCGGTGGCTGTGGCAAGTCAGCCTTGCTA  790
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1301  CATCGTGGTAGGCCGGGACACCAGCTTGGGCCGGTGGCCGTGGCAAGTCAGCCTTGCTA 1360

Query  791   TGATGGAGCACAGCTCTGTGGGGGATCCCTGCTCTCCGGGGACTGGGTGCTGACAGCGC  850
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1361  TGATGGAGCACAGCTCTGTGGGGGATCCCTGCTCTCCGGGGACTGGGTGCTGACAGCGC 1420

Query  851   CCACTGCTTCCCCGAGCGGAACCGGGTCCTGTCCCGATGGCGAGTGTTTGCCGGGTGCCGT  910
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1421  CCACTGCTTCCCGGAGCGGAACCGGGTCCTGTCCCGATGGCGAGTGTTTGCCGGTGCCGT 1480

Query  911   GGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGGGGTGCAGGCTGTGGTCTACCACGGGGG  970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1481  GGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGGGGTGCAGGCTGTGGTCTACCACGGGGG 1540

Query  971   CTATCTTGCCTTTCGGGACCCCAACAGCGAGGAGAACAGCAACGATATTGCCCTGGTCCA 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1541  CTATCTTCCCTTTCGGGACCCCAACAGCGAGGAGAACAGCAACGATATTGCCCTGGTCCA 1600

Query  1031  CCTTTCCAGTCCCCTGCCCCTCACAGAATACATCCAGCCTGTGTGCCTCCCAGCTGCCGG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1601  CCTCTCCAGTCCCCTGCCCCTTCACAGAATACATCCAGCCTGTGTGCCTCCCAGCTGCCGG 1660

Query  1091  GCAGGCCTTGGTGGATGGCAAGATCTGTACCGTGACGGGCTGGGGCAACACGCAGTACTA 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1661  CCAGGCCTTGGTGGATGGCAAGATCTGTACCGTGACGGGCTGGGGCAACACGCAGTACTA 1720

Query  1151  TGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGAGTCCCCATAATCAGCAATGATGTCTG 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1721  TGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGAGTCCCCATAATCAGCAATGATGTCTG 1780

Query  1211  CAATGGGCCTGACTTCTATGGAAACCAGATCAAGCCCAAGATGTTCTGTGCTGGCTACCC 1270
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1791  CAATGGCGCTGACTTTTATGGAAACCAGATCAAGCCCAAGATGTTCTGTGCTGGCTACCC  1850

Query  1071  CGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCTTTGTGTGTGAGGACAG  1130
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1851  CGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCTTTGTGTGTGAGGACAG  1910

Query  1131  CATCTCTCGGACGCCACGTTGGCGGCTGTGTGGCATTGTGAGTTGGGGCACTGGCTGTGC  1190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1911  CATCTCTCGGACGCCACGTTGGCGGCTGTGTGGCATTGTGAGTTGGGGCACTGGCTGTGC  1970

Query  1191  CCTGGCCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGTGGATCTTCCA  1250
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1971  CCTGGCCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGTGGATCTTCCA  2030

Query  1251  GGCCATAAAGACTCACTCCGAAGCCAGCGGCATGGTGACCCAGCTCTGACCGGTGGCTTC  1310
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2031  GGCCATAAAGACTCACTCCGAAGCCAGCGGCATGGTGACCCAGCTCTGACCGGTGGCTTC  2090

Query  1311  TGGCTGCGCAGCCTCCAGGGCCCGAGGTGATCCCGGTGGTGGGATCCACGCTGGGCCGAG  1370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2091  TGGCTGCGCAGCCTCCAGGGCCCGAGGTGATCCCGGTGGTGGGATCCACGCTGGGCCGAG  2150

Query  1371  GATGGGACGTTTTTCTTCTTGGGCCCGGTCCACAGGTCCAAGGACACCCTCCCTCCAGGG  1430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2151  GATGGGACGTTTTTCTTCTTGGGCCCGGTCCACAGGTCCAAGGACACCCTCCCTCCAGGG  2210

Query  1431  TCTTCTGTTCCACAGTGGGGGGGCCACTCAGCCCCGAGACCACCCAACCTCACCCTCTG  1490
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2211  TCTTCTGTTCCACAGTGGGGGGGCCACTCAGCCCCGAGACCACCCAACCTCACCCTCTG  2270

Query  1491  ACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCCCTGATGATGG  1550
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2271  ACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCCCTGATGATGG  2330

Query  1551  GATGCTCTTTAAATAATAAAGATGGTTTTGATT  1583
            ||||||||||||||||||||||||||||||||
Sbjct  2331  GATGCTCTTTAAATAATAAAGATGGTTTTGATT  2363
```



```
 Score =  371 bits (193),  Expect = 4e-99
 Identities = 193/193 (100%), Gaps = 0/193 (0%)
 Strand=Plus/Plus

Query  1    TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGCCAGCCTGGCCTAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGCCAGCCTGGCCTAGC  60

Query  61   AGGCCCCACGCCACCGCCTCTGCCTCCAGCCCGCCCGCTGCTGCGGGGCCACCATGCTCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   AGGCCCCACGCCACCGCCTCTGCCTCCAGCCCGCCCGCTGCTGCGGGGCCACCATGCTCC  120

Query  121  TGCCCAGGCCTGGAGACTGACCCGACCCCGGCACTACCTCGAGGCTCCCGCCCCCACCTGC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  TGCCCAGGCCTGGAGACTGACCCGACCCCGGCACTACCTCGAGGCTCCCGCCCCCACCTGC  180
```

```
Query  181  TGGRCCCCAGGGT  193
            |||||||||||||
Sbjct  181  TGGRCCCCAGGGT  193


CPU time:     0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 529
Number of extensions: 7
Number of successful extensions: 2
Number of sequences better than 10.0: 1
Number of HSP's gapped: 2
Number of HSP's successfully gapped: 2
Length of query: 1793
Length of database: 18,244,403,944
Length adjustment: 27
Effective length of query: 1766
Effective length of database: 18,244,403,917
Effective search space: 32231713278252
Effective search space used: 32231713278252
X1: 11 (21.1 bits)
X2: 26 (50.0 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 6. BLAST alignment of SEQ ID NO:1 with NM_002151.1

Blast 2 Sequences results

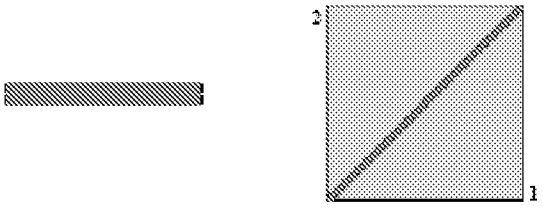
Pubmed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|45644806|ref|NM_002151.1|Homo sapiens hepsin (transmembrane protease, serine 1) (HPN), transcript variant 2, mRNA
Length = 1783 (1 .. 1783)



NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3420 bits (1783),  Expect = 0.0
 Identities = 1783/1783 (100%), Gaps = 0/1783 (0%)
 Strand=Plus/Plus

Query  1     TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60

Query  61    AGGCCCCACGCCACCGCCTCTGCCTCCAGGCCGCCCGCTGCTGCTGGGGGCCACCATGCTCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    AGGCCCCACGCCACCGCCTCTGCCTCCAGGCCGCCCGCTGCTGCTGGGGGCCACCATGCTCC  120

Query  121   TGCCCAGGCCTGGAGACTGACCCACCCCGGCACTACCTCGAGGCTCCGGCCCCCACCTGC  180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1021  CTCTGGTCCACCTCTCCAGTCCCCTGCCCTTCACAGAATACATCCAGCCTGTGTGTCCTCC  1080

Query  1081  CAGCTGCCGGGCCAGGCCCTGGTGGATGGCAAGATCTGTACCGTGACGGGCTGGGGCAACA  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  CAGCTGCCGGGCCAGGCCCTGGTGGATGGCAAGATCTGTACCGTGACGGGCTGGGGCAACA  1140

Query  1141  CGCTACTACTATGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGAGTCCCCATAATCAGCA  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  CGCAGTACTATGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGAGTCCCCATAATCAGCA  1200

Query  1201  ATGATGTCTGCAATGGCGCTGACTTCTATGGAAACCAGATCAAGCCCAAGATGTTCTGTG  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  ATGATGTCTGCAATGGCGCTGACTTCTATGGAAACCAGATCAAGCCCAAGATGTTCTGTG  1260

Query  1261  CTGGCTACCCCGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCTTTGTGT  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  CTGGCTACCCCGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCTTTGTGT  1320

Query  1321  GTGAGGACAGCATCTCTGGGACGCCACGTTGGGGGCTGTGTGGCATTGTGAGTTGGGGCA  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  GTGAGGACAGCATCTCTGGGACGCCACGTTGGGGGCTGTGTGGCATTGTGAGTTGGGGCA  1380

Query  1381  CTGGCTGTGCCCTGGGCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGT  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  CTGGCTGTGCCCTGGGCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGT  1440

Query  1441  GGATCTTCCAGGCCATAAAGACTCACTCCGAAGCCAGCGGCATGCTGACCCAGCTCTGAC  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  GGATCTTCCAGGCCATAAAGACTCACTCCGAAGCCAGCGGCATGCTGACCCAGCTCTGAC  1500

Query  1501  CGGTGGCTTCTCTGGCTGGGCAGCGTGCAGGGCCCGAGGTGATCCGGTGGTGGGATCCACG  1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  CGGTGGCTTCTCTGGCTGCGCAGCCTCCAGGGCCCGAGGTGATCCCGGTGGTGGGATCCACG  1560

Query  1561  CTGGGCCGAGGATGTGACTTTTTCTTCTTGGGCCCGGTCCACAGTCCGAGGACACCCT  1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  CTGGGCCGAGGATGTGGACGTTTTCTTCTTGGGCCCGGTCCACAGGTCCAAGGACACCCT  1620

Query  1621  CCCTCCAGGGTGCTCTCTTCCACAGTGGCGGGCCGACTGAGCCCCGAGACCACCCAACCT  1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  CCCTCCAGGGTGCTCTCTTCCACAGTGGCGGGCCGACTGAGCCCCGAGACCACCCAACCT  1680

Query  1681  CACCCTCTTGACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCC  1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  CACCCTCCTGACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCC  1740

Query  1741  CTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
              |||||||||||||||||||||||||||||||||||||||||||
Sbjct  1741  CTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
```

```
CPU time:     0.02 user secs.       0.01 sys. secs          0.03 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H

    1.33    0.621     1.12


 Matrix: blastn matrix:1 -2
 Gap Penalties: Existence: 5, Extension: 2
 Number of Sequences: 1
 Number of Hits to DB: 505
 Number of extensions: 5
 Number of successful extensions: 1
 Number of sequences better than 10.0: 1
 Number of HSP's gapped: 1
 Number of HSP's successfully gapped: 1
 Length of query: 1733
 Length of database: 13,244,403,944
 Length adjustment: 27
 Effective length of query: 1706
 Effective length of database: 13,244,403,917
 Effective search space: 22593178278262
 Effective search space used: 22593178278262
 X1: 11 (21.1 bits)
 X2: 28 (50.0 bits)
 X3: 26 (50.0 bits)
 S1: 14 (27.6 bits)
 S2: 22 (43.0 bits)
```

Exhibit 7. BLAST alignment of SEQ ID NO:1 with X07732.1

Blast 2 Sequences results

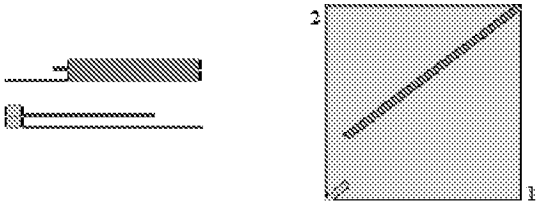
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 0.0001 wordsize: 11 Filter: View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

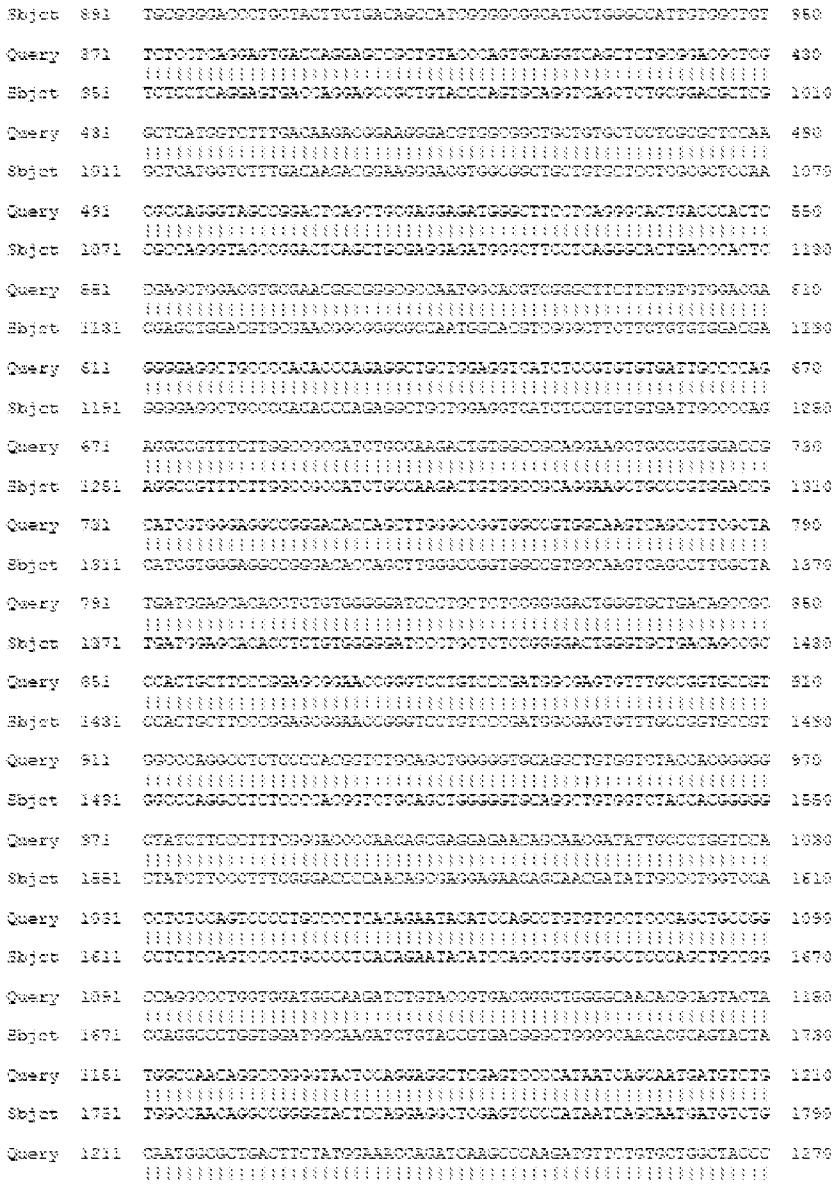
Sequence 2: gi|32363|emb|X07732.1|Human hepatoma mRNA for serine protease hepsin
Length = 2363 (1 .. 2363)



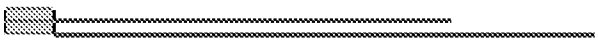
NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3063 bits (1593),  Expect = 0.0
 Identities = 1593/1593 (100%), Gaps = 0/1593 (0%)
 Strand=Plus/Plus

Query  191   GGTCCCACCCTGGCCCAGGAGGTCAGTCAGGGAATCATTAACAAGAGGCAGTGACATGCC  250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  771   GGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATCATTAACAAGAGGCAGTGACATGGC  830

Query  251   GCAGAAGGAGGGTGGCCGGACTGTGCCATGCTGCTCCAGACCCAAGGTGGCAGGTGTCAC  310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  831   GCAGAAGGAGGGTGGCCGGACTTTGCCATGCTGCTCCAGACCCAAGGTGGCAGCTCTCAC  890

Query  311   TGCGGGGACCCTGCTACTTCTGACAGCCATCGGGGCGGCATCCTGGCCATTGTGGCTGT  370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



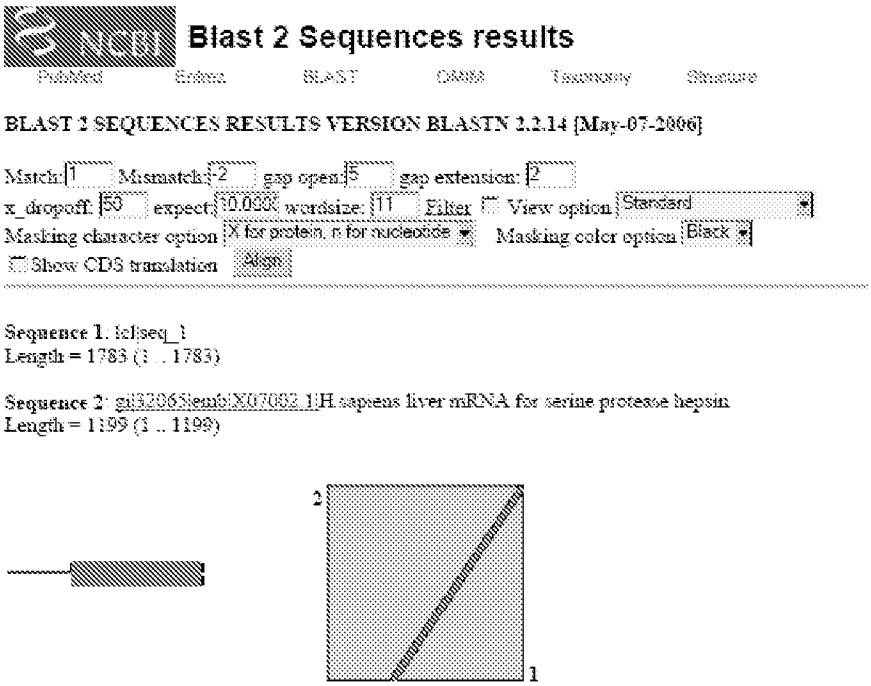
```
Sbjct  1781  CAATGGCGCTGACTTCTATGGAAACCAGATCAAGCCCAAGATGTTCTGTGCTGGCTACCC  1840

Query  1271  CGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCCTTGTGTGTGAGGACAG  1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1841  CGAGGGTGGCATTGATGCCTGCCAGGGCGACAGCGGTGGTCCCCTTGTGTGTGAGGACAG  1900

Query  1331  CATCTCTCGGACGCCACGTTGGCGGCTGTGTGGCATTGTGAGTTGGGGCACTGGCTGTGC  1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1901  CATCTCTCGGACGCCACGTTGGCGGCTGTGTGGCATTGTGAGTTGGGGCACTGGCTGTGC  1960

Query  1391  CCTGGCCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGTGGATCTTCCA  1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1961  CCTGGCCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCGGGAGTGGATCTTCCA  2020

Query  1451  GGCCATAAAGACTCACTCCGAAGCCAGCGGCATGGTGACCCAGCTCTGACCGGTGGCTTC  1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2021  GGCCATAAAGACTCACTCCGAAGCCAGCGGCATGGTGACCCAGCTCTGACCGGTGGCTTC  2080

Query  1511  TCGCTGGGCAGCCTCCAGGGCCCGAGGTGATCCCGGTGGTGGGATCCACGCTGGGCCGAG  1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2081  TCGCTGCGCAGCTTCCAGGGCCCGAGGTGATCCCGGTGGTGGGATCCACGCTGGGCCGAG  2140

Query  1571  GATGGGACGTTTTTCTTCTTGGGCCTGGTCCACAGGTCCAAGGACACCTTCCCTCCAGGG  1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2141  GATGGGACGTTTTTCTTCTTGGGCCCGGTCCACAGGTCCAAGGACACCTTCCCTCCAGGG  2200

Query  1631  TCCTCTCTTTCCACAGTGGCGGGCCCACTCAGCCCCGAGACCACCCAACCTCACCCTCCTG  1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2201  TCCTCTCTTTCCACAGTGGCGGGCCCACTCAGCCCCGAGACCACCCAACCTCACCCTCCTG  2260

Query  1691  ACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCCCTGATGATGG  1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2261  ACCCCCATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTAGGTGCCCCTGATGATGG  2320

Query  1751  GATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
              |||||||||||||||||||||||||||||||||
Sbjct  2321  GATGCTCTTTAAATAATAAAGATGGTTTTGATT  2353
```



```
 Score =  871 bits (159),  Expect = 4e-99
 Identities = 159/159 (100%), Gaps = 0/159 (0%)
 Strand=Plus/Plus

Query  1    TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    TCGAGCCCGCTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60

Query  61   AGGCCCCACGCCACCGCCTCTGCCTCTAGGCCGCCCGCTGCTGCGGGCCTACCATGCTCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   AGGCCCCACGCCACTGCCTCTGCCTCCAGGCCGCCCGCTGCTGCGGGCCCACCATGCTCC  119

Query  121  TGCCCAGGCCTGGAGACTGACCCGACCCCGGCACTACCTCGAGGCTCCGCCCCCACCTGC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  TGCCCAGGCCTGGAGACTGACCCGACCCCGGCACTACCTCGAGGCTCCGCCCCCACCTGC  180
```

```
Query  181  TGGACCCCAGGGT  193
            |||||||||||||
Sbjct  181  TGGACCCCAGGGT  193


CPU time:     0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 520
Number of extensions: 7
Number of successful extensions: 2
Number of sequences better than 10.0: 1
Number of HSP's gapped: 2
Number of HSP's successfully gapped: 2
Length of query: 1783
Length of database: 15,244,403,944
Length adjustment: 27
Effective length of query: 1756
Effective length of database: 15,244,403,917
Effective search space: 26767179278252
Effective search space used: 26767179278252
X1: 11 (21.1 bits)
X2: 25 (50.0 bits)
X3: 25 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 8. BLAST alignment of SEQ ID NO:1 with X07002.1



NCBI **Blast 2 Sequences results**

Publmed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 0.0002 wordsize: 11 Filter View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|3206|emb|X07002.1|H.sapeus liver mRNA for serine protease hepsin
Length = 1199 (1 .. 1199)

NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 2305 bits (1199),  Expect = 0.0
 Identities = 1199/1199 (100%), Gaps = 0/1199 (0%)
 Strand=Plus/Plus

Query  585   ACGTCGGGCTTCTTCTGTGTGTGGACGAGGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAG  644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1      ACGTCGGGCTTCTTCTGTGTGGACGAGGGGAGGCTGCCCCACACCCAGAGGCTGCTGGAG  60

Query  645   GTCATCTCCGTGTGTGATTGCCCCAGAGGCCGTTTCTTGGCCGCCATCTGCCAAGACTGT  704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GTCATCTCCGTGTGTGATTGCCCCAGAGGCCGTTTCTTGGCCGCCATCTGCCAAGACTGT  120

Query  705   GGCCGCAGGAAGCTGCCCGTGGACCGCATCGTGGGAGGCCCGGACGCCGCTTGGGCCGG  764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  121   GGCCGCAGGAAGTTGCCCGTGGACCGCATCGTGGGAGGCCGGGACACCAGCTTGGGCCGG  180

Query  768   TGGCCGTGGCAAGTCAGCCTTCGCTATGATGGAGCACACTTCTGTGGGGGATCCCTGCTC  827
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181   TGGCCGTGGCAAGTCAGCCTTCGCTATGATGGAGCACACCTCTGTGGGGGATCCCTGCTC  240

Query  828   TCCGGGGACTGGGTGCTGACAGCCGCCCACTGCTTCCCGGAGCGGAACCGGGTCCTGTCC  887
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241   TCTGGGGACTGGGTGCTGACAGCCGCCCACTGCTTCCCGGAGCGGAACCGGGTCCTGTCC  300

Query  888   CGATGGCGAGTGTTTGCCGGTGCCGTGGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGG  947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301   CGATGGCGAGTGTTTGCCGGTGCCGTGGCCCAGGCCTCTCCCCACGGTCTGCAGCTGGGG  360

Query  948   GTGCAGGCTGTGGTCTACCACGGGGGCTATCTTCCCTTTCGGGACCCCAACAGCGAGGAG  1007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361   GTGCAGGCTGTGGTCTACCACGGGGGCTATCTTCCCTTTCGGGACCCCAACAGCGAGGAG  420

Query  1008  AACAGCAACGATATTGCCCTGGTCCAGCTCTCCAGTCCCCTGCCCCTCACAGAATACATC  1067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421   AACAGCAACGATATTGCCCTGGTCCAGCTCTCCAGTCCCCTGCCCCTCACAGAATACATC  480

Query  1068  CAGCCTGTGTGCCTCCCAGCTGCCGGCTAGGCCCTGGTGGATGGCAAGATCTGTACCGTG  1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481   CAGCCTGTGTGCCTCCCAGCTGCCGGCCAGGCCCTGGTGGATGGCAAGATCTGTACCGTG  540

Query  1128  ACGGGCTGGGGCAACAGCAGCAGTACTATGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGA  1187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541   ACGGGCTGGGGCAACAGCAGCAGTACTATGGCCAACAGGCCGGGGTACTCCAGGAGGCTCGA  600

Query  1188  GTCCCCATAATCAGCAATGATGTCTGCAATGGCGCTGACTTCTATGGAAAGCAGATCAAG  1247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601   GTCCCCATAATCAGCAATGATGTCTGCAATGGCGCTGACTTCTATGGAAAGCAGATCAAG  660

Query  1248  CCCAAGATGTTCTGTGCTGGCIACCCCGAGGGTGGCATTGATGCCTGCCAGGGCGACAGC  1307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661   CCCAAGATGTTCTGTGCTGGCIACCCCGAGGGTGGCATTGATGCCTGCCAGGGCGACAGC  720

Query  1308  GGTGGTCCCTTTGTGTGTGAGGACAGCATCTCTGGAAGGCCAGTTGGCGGCTGTGTGGC  1367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721   GGTGGTCCCTTTGTGTGTGAGGACAGCATCTCTGGAAGGCCAGTTGGCGGCTGTGTGGC  780

Query  1368  ATTGTGAGTTGGGGCACTGGCTGTGCCTTGGCCAGAAGCCAGGCGTCTACACCAAAGTC  1427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781   ATTGTGAGTTGGGGCACTGGCTGTGCCTTGGCCAGAAGCCAGGCGTCTACACCAAAGTC  840

Query  1428  AGTGACTTCCGGGAGTGGATGTTCCAGGCCATAAAGACTCACTCCGAAGCCAGCGGCATG  1487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841   AGTGACTTCCGGGAGTGGATGTTCCAGGCCATAAAGACTCACTCCGAAGCCAGCGGCATG  900

Query  1488  GTGACCCAGCTCTGAACGGGTGGCTTCTGCTGCGCAGCTTCCAGGGCCCGAGGTGATCCC  1547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901   GTGACCCAGCTCTGAACGGGTGGCTTCTGCTGCGCAGCCTCCAGGGCCCGAGGTGATCCC  960

Query  1548  GGTGGTGGGATCCACGCTGGGCCGAGGATGGGAGGTTTTCTTCTTGGGCCCGGTCCACA  1607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961   GGTGGTGGGATCCACGCTGGGCCGAGGATGGGAGGTTTTCTTCTTGGGCCCGGTCCACA  1020

Query  1608  GGTCCAAGGACACCCTCCCTCCAGGGTCCTCTTCTTCCACAGTGGCGGGCCCACTCAGCCC  1667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1021  GGTGCAAGGACACCCTCCCTCCAGGGTCCTCTCTTCCACAGTGGCGGGGCCACTCAGGCC  1080

Query  1665  CGAGACCACCCAACCTCACCCTCCTGACCCCCATGTAAATATTGTTCTGCTGTCTGGGAC  1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  CGAGACCACCCAACCTCACCCTCCTGACCCCCATGTAAATATTGTTCTGCTGTCTGGGAC  1140

Query  1725  TCCTGTCTAGGTGCCCCTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  TCCTGTCTAGGTGCCCCTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1199


CPU time:     0.02 user secs.       0.01 sys. secs          0.03 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 826
Number of extensions: 9
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1783
Length of database: 18,244,403,944
Length adjustment: 27
Effective length of query: 1756
Effective length of database: 18,244,403,917
Effective search space: 32037178278252
Effective search space used: 32037178278252
X1: 11 (21.1 bits)
X2: 28 (50.5 bits)
X3: 28 (50.5 bits)
S1: 18 (32.7 bits)
S2: 22 (40.0 bits)
```

Exhibit 9. BLAST alignment of SEQ ID NO:1 with rat and mouse hepsin coding sequences

NCBI **Blast 2 Sequences results**

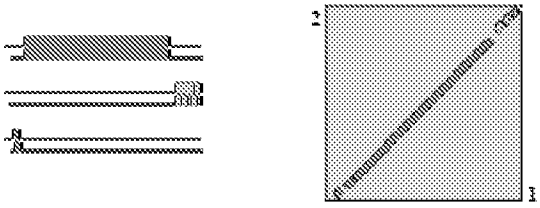
Protein Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 0.005 wordsize: 11 Filter View option Standard
Masking character option X for protein, n for nucleotide Masking color option Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|37928|emb|X70900.1|R.norvegicus mRNA for hepsin
Length = 1739 (1 .. 1739)

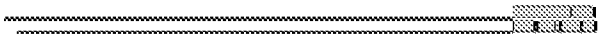


NOTE: Bitscore and expect value are calculated based on the size of the nr database.

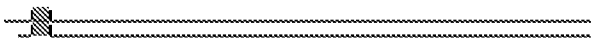
NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 1908 bits (680),  Expect = 0.0
 Identities = 1107/1318 (83%), Gaps = 8/1318 (0%)
 Strand=Plus/Plus

Query  184   ACCCCAGGGTCCCACCCTGGCCCAGGAGGTCAGCCAGGGAATCATTAACAAGAGGCAGTG  243
             ||||||||||  ||  |||   |||||  ||||||  ||  ||||||||||||||||||  |   ||
Sbjct  128   ACCCCAGGGTTCCGCCCCAGCCCAACAGGTCAACCTGGGAATCATTAACAAGAGTCCCTG  187

Query  244   ACATGGCGCAGAAGGAGGGTGGCCCGACTGTGCCATGCTGCTCCAGACCCAAGGTGGCAG  303
             ||||||||   ||||||||||||||||||||  ||||||||  ||||||||||||||||||||
Sbjct  188   ACATGGCG---AAGGAGGGTGGCCGGACTGCACCATGCTGTTCCAGACCCAAGGTGGCAG  244

Query  304   CTCTCACTGCGGGGGACGCTGCTACTTCTGAACAGCCATCGGGGCGGCATCCTGGGCCATTG  363
             |||||||||  ||||||||||||  | |||||||  ||| |||||  || ||||||||||||
```

```
Sbjct  1140  AAGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAACCCAAGATGTTCTGTGCTG  1195

Query  1264  GCTACCCCGAGGGTGGCATTGATGCTTGTCAGGGCGACAGCGGTGGTCCCTTTGTGTGTG  1323
              !!!! !! !!!!!!!! !!!!!!!! !!!!!!!! !!!!!!! !! ! !!!!!! !!!!
Sbjct  1200  GCTATCCTGAGGTTGGTATTGATGCATGCCAGGGTGACAGTGGAGGCCACTTTGTATGTG  1259

Query  1324  AGGACAGCATCTCTGGAACATCAACATGGCGGCTCTGGCGGCATTGTGAAGTTGGGGGACTG  1383
              !!!!!!! !!!!!! ! !!  !! ! !!!!!!!! !! !!!!!!!! !! !!!!! !! !
Sbjct  1260  AGGACAGAATCTCTGGAACATCAACATGGCGGCTCTGGCGGCATTGTAAGCTGGGGTACCG  1319

Query  1384  GCTGTGCCCTGGCCCAGAAGCCAGGCGTCTACACCAAAGTCAGTGACTTCCCGGGAGTGGA  1443
              !!!!!!! !!!!!! !!!!!! !! !! !!!!!!!!!!!!!! !!!!!!!!!!!!!!!!
Sbjct  1320  GCTGTGCTTTGGCCCGGAAGCCGGGAGTGTACACCAAAGTCATTGACTTCCGGGAGTGGA  1379

Query  1444  TCTTCCAGGCCATAAAGACTCACTTCCGAAGCCAGGGCATGGTGACCCAGCTCTGACC  1501
              !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!! ! !!!!!!!!! !! !!!! !!!!!!
Sbjct  1380  TCTTCCAGGCCATAAAGACTCACTCCGAAGCTACCGGCATGGTAACTCAGCCCTGACC  1437
```



```
 Score = 77.6 bits (40),  Expect = 1e-10
 Identities = 175/230 (76%), Gaps = 9/230 (3%)
 Strand=Plus/Plus

Query  1585  TCCACGCTGGGCCGAGGATGGGACGTTTTTCTTCTTGGGCCCGGTCCACAGGTCCAAGGA  1614
              !!!!! !!!!!!!    !!!! !!! !!!!!! !! !!  !! !!!!! !! !!!!!!!!
Sbjct  1317  TCCACACTGGGGCCGACATGGAACGGTTTTCTGCTCGGATCCAGTCCATAGATCCAAGGA  1376

Query  1615  CAGCCGTCCCTCCAGGGTCCTCTCTTCCACAGTGGCCGGCCCACTCAGCCCCAAGACCACC  1674
                !  !   !!!! !! !!!!!!!!!!!!!!!!!!! !!!!!!!!!!  !!!  ! !!!
Sbjct  1377  TGC--CGGGTCTAAGGACCTCTCTTCCACAGTGGCCGGCCCACTCAATCCCAGGGCCATT  1434

Query  1675  CAACCTCAGCCTCCTGACCGCCATGTAAATATTGTTCCTGCTGTCTGGGA-CTCCTGTCTA  1732
                !!!!!!!!!!!!    !!!!!!!!!!!!!!!!  !!!!   !!!!!!  !! !! !! !
Sbjct  1435  GG-CCCTAGCCTCCCA--CCCCATGTAAATATTACTCCTGTCCTCTGGGGGCTGCTTTCCA  1491

Query  1733  GGTGCCCCTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
              !! !!!!!!  !!   !!!!!!!!!!!!!!!!!!!!!!!! !!!!!!!!!!!
Sbjct  1492  GGCGCCCCTTGTGC--GGATGCTCTTTAAATAATAAAGGTGGTTTTGATT  1539
```



```
 Score = 66.1 bits (34),  Expect = 4e-07
 Identities = 48/55 (87%), Gaps = 0/55 (0%)
 Strand=Plus/Plus

Query  89   GGCCGCCCGCTGCTGCGGGGGCCACCATGCTCCTGCCCAGGCCTGGAGACTGACCC  143
            !!! !!!!!!!!!! !!!! !!!!!!!! !!! !!!!!!!! !!!!!!! !!!!
Sbjct  45   GGCTGCCCGCTGCTGTGGGGACACCATGCCCCTTCCCAGGCCCGGAGACTAACCC  99
```

```
CPU time:     0.03 user secs.       0.00 sys. secs          0.03 total secs.

Lambda      K      H
    1.33    0.621     1.12

Gapped
Lambda      K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 147
Number of extensions: 9
Number of successful extensions: 6
Number of sequences better than 10.0: 1
Number of HSP's gapped: 3
Number of HSP's successfully gapped: 3
Length of query: 1783
Length of database: 18,244,403,344
Length adjustment: 27
Effective length of query: 1756
Effective length of database: 18,244,403,317
Effective search space: 32037732278252
Effective search space used: 32037732278252
X1: 11 (21.1 bits)
X2: 26 (50.0 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.8 bits)
```